

Robust Kantorovich's theorem on Newton's method under majorant condition in Riemannian Manifolds

T. Bittencourt *

O. P. Ferreira[†]

May 19, 2015

Abstract

A robust affine invariant version of Kantorovich's theorem on Newton's method, for finding a zero of a differentiable vector field defined on a complete Riemannian manifold, is presented in this paper. In the analysis presented, the classical Lipschitz condition is relaxed by using a general majorant function, which allow to establish existence and local uniqueness of the solution as well as unifying previously results pertaining Newton's method. The most important in our analysis is the robustness, namely, is given a prescribed ball, around the point satisfying Kantorovich's assumptions, ensuring convergence of the method for any starting point in this ball. Moreover, bounds for Q -quadratic convergence of the method which depend on the majorant function is obtained.

Keywords: Newton's method, robust Kantorovich's theorem, majorant function, vector field, Riemannian manifold

1 Introduction

Extension of concepts and techniques as well as methods of Mathematical Programming from the Euclidean space to Riemannian setting it is natural and has been done frequently before; see, e.g., [1, 2, 4, 15, 27, 38, 44]. The motivation of this extensions, which in general is nontrivial, is either of purely theoretical nature or aims at obtaining efficient algorithms; see, e.g., [1, 2, 12, 21, 32, 27, 28, 38, 44]. Indeed, many optimization problems are naturally posed on Riemannian manifolds, which has a specific underlying geometric and algebraic structure that could be exploited to greatly reduce the cost of obtaining the solutions. For instance, in order to take advantage of the Riemannian geometric structure, it is suitable to treat some constrained optimization problems as one of finding the zeros of a gradient vector field on a Riemannian manifolds rather than use the method of Lagrange multipliers or projection idea for solving the problem; see [1, 2, 27, 38, 44]. In this case, constrained optimization problems can be seen as unconstrained one from the Riemannian geometry viewpoint. Besides, the Riemannian geometry allows to induce new research directions so as to produce competitive algorithms; see [12, 21, 32, 38]. In this paper, instead of considering the problem of finding the zero of the gradient field on a Riemannian manifolds, let us consider the more general problem of finding a zeros of a vector field defined on a Riemannian manifold.

*IME/UFG, CP-131, CEP 74001-970 - Goiânia, GO, Brazil (Email: tiberio.b@gmail.com). This author was supported by CAPES.

[†]IME/UFG, CP-131, CEP 74001-970 - Goiânia, GO, Brazil (Email: orizon@mat.ufg.br). This author was partially supported by FAPEG 201210267000909 - 05/2012, PRONEX-Optimization(FAPERJ/CNPq), CNPq Grants 4471815/2012-8, 305158/2014-7.

On the other hand, the Newton's method and its variant are powerful tools for finding a zero of nonlinear function in real or complex Banach space. Besides its practical applications, Newton's method is also a powerful theoretical tool having a wide range of applications in pure mathematics; see [8, 19, 22, 29, 30, 43]. Therefore, a couple of papers have dealt with the issue of generalization of Newton's method and its variant from Euclidean to Riemannian setting in order to go further in the study of the convergence properties of this method. Early works dealing with the generalization of Newton's methods to Riemannian setting include [10, 12, 18, 33, 36, 39]. Actually, the generalization of Newton's method to Riemannian setting has been done with several different purposes, including the purpose of finding a zeros of a gradient vector field or, more generally, with the purpose of finding a zero of a differentiable vector field; see [1, 2, 5, 6, 7, 9, 10, 12, 14, 15, 24, 25, 26, 27, 28, 34, 35, 38, 40, 41, 42, 46] and the references therein.

Properties of convergence of Newton's method have been extensively studied on several papers due to the important role that it plays in the development of numerical methods for finding a zero of a differentiable vector field defined on a complete Riemannian manifolds. In 2002 Ferreira and Svaiter in [15] extended the Kantorovich's theorem on the Newton's method to Riemannian setting using a new technique which simplifies the analysis and proof of this theorem. It is worth mention that, in a similar spirit, an extensions of the famous Smale's theory; see [37], to analytic vector fields on analytic Riemannian manifolds were done in 2003 by Dedieu et al. in [9]. The basic idea of [15] was to combine a formulation of Kantorovich's theorem by means of quadratic majorant functions, see [45] for more general majorant functions, with the definitions of good regions for the Newton's method. In these regions, the majorant function bounds the vector fields which the zero is to be found, and the behavior of the Newton's iteration in these regions is estimated using iterations associated to the majorant function. Moreover, as a whole, the union of all these regions is invariant under Newton's iteration. Afterward, this technique was successfully employed for proving generalized versions of Kantorovich's theorem in Riemannian setting. Inspired by previous work of Zabrejko and Nguen in [45] on Kantorovich's majorant method, a radial parametrization of a Lipschitz-type and L-average Lipschitz affine invariant majorante conditions were introduced in Riemannian setting by Alvarez et al. in [3] and Li and Wang in [25], respectively, in order to establish existence and local uniqueness of the solution as well as unifying previously convergence criterion of Newton's method.

In the present paper, we will use the technique introduced in [15], see also [17], to present a robust affine invariant version of the Kantorovich's theorem on the Newton's method finding a zeros of a differentiable vector field defined on a complete Riemannian manifold. In our analysis, the classical Lipschitz condition is relaxed using a general majorant function. The analysis presented provides a clear relationship between the majorant function and the vector field under consideration. However, the most important in our analysis is the robustness, namely, we give a prescribed ball, around the point satisfying the Kantorovich's assumptions, ensuring convergence of the method for any starting point in this ball. Moreover, we establish bounds for Q -quadratic convergence of the method which depend on the majorant function. Also, as in [3] and [25], this analysis allows us establish existence and local uniqueness of the solution as well as unifying previously results pertaining Newton's method.

The organization of the paper is as follows. In Section 1.1, some notations and one basic results used in the paper are presented. In Section 2, the main result is stated, namely, the robust affine invariant Kantorovich's theorem for Newton's method and in Section 2 the affine invariant version, which is used for proving the robust one is stated and proved. In Section 4 we prove the main theorem. In Section 5 three special case of the main theorem is presented. Some final remarks are made in Section 6.

1.1 Notation and auxiliary results

In this section we recall some notations, definitions and basic properties of Riemannian manifolds used throughout the paper, they can be found, for example, in [11] and [23].

Throughout the paper, \mathcal{M} is a smooth manifold and $C^1(\mathcal{M})$ is the class of all continuously differentiable functions on \mathcal{M} . The space of vector fields $C^r(\mathcal{M})$ on \mathcal{M} is denoted by $\mathcal{X}^r(\mathcal{M})$, by $T_p\mathcal{M}$ we denote the tangent space of \mathcal{M} at p and by $T\mathcal{M} = \bigcup_{x \in \mathcal{M}} T_x\mathcal{M}$ the *tangent bundle* of \mathcal{M} . Let \mathcal{M} be endowed with a Riemannian metric $\langle \cdot, \cdot \rangle$, with corresponding norm denoted by $\|\cdot\|$, so that \mathcal{M} is now a *Riemannian manifold*. Let us recall that the metric can be used to define the length of a piecewise C^1 curve $\zeta : [a, b] \rightarrow \mathcal{M}$ by

$$\ell[\zeta, a, b] := \int_a^b \|\zeta'(t)\| dt.$$

Minimizing this length functional over the set of all such curves we obtain a distance $d(p, q)$, which induces the original topology on \mathcal{M} . The open and closed balls of radius $r > 0$ centered at p are defined, respectively, as

$$B(p, r) := \{q \in \mathcal{M} : d(p, q) < r\}, \quad B[p, r] := \{q \in \mathcal{M} : d(p, q) \leq r\}.$$

Let ζ be a curve joining the points p and q in \mathcal{M} and let ∇ be the Levi-Civita connection associated to $(\mathcal{M}, \langle \cdot, \cdot \rangle)$. For each $t \in [a, b]$, ∇ induces an isometry, relative to $\langle \cdot, \cdot \rangle$,

$$\begin{aligned} P_{\zeta, a, t} : T_{\zeta(a)}\mathcal{M} &\longrightarrow T_{\zeta(t)}\mathcal{M} \\ v &\longmapsto P_{\zeta, a, t}v = V(t), \end{aligned} \tag{1}$$

where V is the unique vector field on ζ such that $\nabla_{\zeta'(t)}V(t) = 0$ and $V(a) = v$, the so-called *parallel transport* along ζ from $\zeta(a)$ to $\zeta(t)$. Note also that

$$P_{\zeta, b_1, b_2} \circ P_{\zeta, a, b_1} = P_{\zeta, a, b_2}, \quad P_{\zeta, b, a} = P_{\zeta, a, b}^{-1}. \tag{2}$$

A vector field V along ζ is said to be *parallel* if $\nabla_{\zeta'}V = 0$. If ζ' itself is parallel, then we say that ζ is a *geodesic*. The geodesic equation $\nabla_{\zeta'}\zeta' = 0$ is a second order nonlinear ordinary differential equation, so the geodesic ζ is determined by its position p and velocity v at p . The restriction of a geodesic to a closed bounded interval is called a *geodesic segment*. It is easy to check that $\|\zeta'\|$ is constant. We usually do not distinguish between a geodesic and its geodesic segment, as no confusion can arise. We say that ζ is *normalized* if $\|\zeta'\| = 1$. A geodesic $\zeta : [a, b] \rightarrow \mathcal{M}$ is said to be *minimal* if its length is equal the distance of its end points, i.e. $\ell[\zeta, a, b] = d(\zeta(a), \zeta(b))$.

A Riemannian manifold is *complete* if its geodesics are defined for any values of t . *In this paper, all manifolds \mathcal{M} are assumed to be complete.* The Hopf-Rinow's theorem asserts that if this is the case then any pair of points, say p and q , in \mathcal{M} can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, (\mathcal{M}, d) is a complete metric space and bounded and closed subsets are compact. The *exponential map* at p , $\exp_p : T_p\mathcal{M} \rightarrow \mathcal{M}$, is defined by $\exp_p v = \zeta_v(1)$, where ζ_v is the geodesic defined by its position p and velocity v at p and $\zeta_v(t) = \exp_p tv$ for any value of t .

Let $X \in C^1(\mathcal{M})$. The covariant derivative of X determined by the Levi-Civita connection ∇ defines at each $p \in \mathcal{M}$ a linear map $\nabla X(p) : T_p\mathcal{M} \rightarrow T_p\mathcal{M}$ given by

$$\nabla X(p)v := \nabla_Y X(p), \tag{3}$$

where Y is a vector field such that $Y(p) = v$.

Definition 1. Let Y_1, \dots, Y_n be vector fields on \mathcal{M} . Then, the n -th covariant derivative of X with respect to Y_1, \dots, Y_n is defined inductively by

$$\nabla_{\{Y_1, Y_2\}}^2 X := \nabla_{Y_2} \nabla_{Y_1} X, \quad \nabla_{\{Y_i\}_{i=1}^n}^n X := \nabla_{Y_n} (\nabla_{Y_{n-1}} \cdots \nabla_{Y_1} X).$$

Definition 2. Let $p \in \mathcal{M}$. Then, the n -th covariant derivative of X at p is the n -th multilinear map $\nabla^n X(p) : T_p \mathcal{M} \times \dots \times T_p \mathcal{M} \rightarrow T_p \mathcal{M}$ defined by

$$\nabla^n X(p)(v_1, \dots, v_n) := \nabla_{\{Y_i\}_{i=1}^n}^n X(p),$$

where Y_1, \dots, Y_n are vector fields on \mathcal{M} such that $Y_1(p) = v_1, \dots, Y_n(p) = v_n$.

We remark that Definition 2 only depends on the n -tuple of vectors (v_1, \dots, v_n) since the covariant derivative is tensorial in each vector field Y_i .

Definition 3. Let $p \in \mathcal{M}$. The norm of an n -th multilinear map $A : T_p \mathcal{M} \times \dots \times T_p \mathcal{M} \rightarrow T_p \mathcal{M}$ is defined by

$$\|A\| = \sup \{ \|A(v_1, \dots, v_n)\| : v_1, \dots, v_n \in T_p \mathcal{M}, \|v_i\| = 1, i = 1, \dots, n \}.$$

In particular, the norm of the n -th covariant derivative of X at p is given by

$$\|\nabla^n X(p)\| = \sup \{ \|\nabla^n X(p)(v_1, \dots, v_n)\| : v_1, \dots, v_n \in T_p \mathcal{M}, \|v_i\| = 1, i = 1, \dots, n \}.$$

Now, the *Fundamental Theorem of Calculus* for a vector field X becomes

Lemma 1. Let Ω be an open subset of \mathcal{M} , X a C^1 vector field defined on Ω and $\zeta : [a, b] \rightarrow \Omega$ a C^1 curve. Then

$$P_{\zeta, t, a} X(\zeta(t)) = X(\zeta(a)) + \int_a^t P_{\zeta, s, a} \nabla X(\zeta(s)) \zeta'(s) ds, \quad t \in [a, b].$$

Proof. See [15]. □

Lemma 2. Let Ω be an open subset of \mathcal{M} , X a C^2 vector field defined on Ω and $\zeta : [a, b] \rightarrow \Omega$ a C^1 curve. Then for all $Y \in \mathcal{X}(\mathcal{M})$ we have

$$P_{\zeta, t, a} \nabla X(\zeta(t)) Y(\zeta(t)) = \nabla X(\zeta(a)) Y(\zeta(a)) + \int_a^t P_{\zeta, s, a} \nabla^2 X(\zeta(s)) (Y(\zeta(s)), \zeta'(s)) ds, \quad t \in [a, b].$$

Proof. See [24]. □

Lemma 3 (Banach's Lemma). Let B be a linear operator and let I_p be the identity operator in $T_p \mathcal{M}$. If $\|B - I_p\| < 1$ then B is nonsingular and $\|B^{-1}\| \leq 1 / (1 - \|B - I_p\|)$.

Proof. See, for example, [37]. □

We also need the following elementary convex analysis result, see [20]:

Proposition 1. Let $I \subset \mathbb{R}$ be an interval and $\varphi : I \rightarrow \mathbb{R}$ be convex. For any $s_0 \in \text{int}(I)$, the left derivative there exist (in \mathbb{R})

$$D^- \varphi(s_0) := \lim_{s \rightarrow s_0^-} \frac{\varphi(s_0) - \varphi(s)}{s_0 - s} = \sup_{s < s_0} \frac{\varphi(s_0) - \varphi(s)}{s_0 - s}.$$

Moreover, if $s, t, r \in I$, $s < r$, and $s \leq t \leq r$ then $\varphi(t) - \varphi(s) \leq [\varphi(r) - \varphi(s)] [(t - s)/(r - s)]$.

2 Robust Kantorovich's Theorem on Newton's Method

Our goal is to state and prove a robust affine invariant version of Kantorovich's Theorem on Newton's Method for finding a zero of a vector field:

$$X(p) = 0, \quad (4)$$

where \mathcal{M} is a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \rightarrow T\mathcal{M}$ a continuously differentiable vector field. The most important in our analysis is the robustness, namely, we give a prescribed ball, around the point satisfying the Kantorovich's assumptions, ensuring convergence of the method for any starting point in this ball. Moreover, we establish bounds for Q -quadratic convergence of the method which depend on the majorant function. Also, as in [3] and [25], this analysis allows us establish existence and local uniqueness of the solution. For state the theorem we need some definitions. We beginning with the following definition which was introduced in [3].

Definition 4. Let $R > 0$, $n \in \mathbb{N} \setminus \{0\}$, $p_0 \in \mathcal{M}$ and $\mathcal{G}_n(p_0, R)$ be the class of all piecewise geodesic curves $\xi : [0, T] \rightarrow \mathcal{M}$ for some $T > 0$ which satisfy the following conditions:

1. $\xi(0) = p_0$ and the length of ξ is no greater than R ;
2. there exist $c_0, c_1, \dots, c_n \in [0, T]$ with $c_0 = 0 \leq c_1 \leq \dots \leq c_n = T$ such that $\xi|_{[c_0, c_1]}$, $\dots, \xi|_{[c_{n-2}, c_{n-1}]}$ are $n-1$ minimizing geodesics and $\xi|_{[c_{n-1}, c_n]}$ is a geodesic.

Remark 1. Since \mathcal{M} is complete, $\mathcal{G}_n(p_0, R)$ is nonempty. Moreover, $\mathcal{G}_n(p_0, R) \subset \mathcal{G}_{n+1}(p_0, R)$ for all $n \in \mathbb{N} \setminus \{0\}$. Note that, in Definition 4, $\mathcal{G}_1(p_0, R)$ is the class of all minimizing geodesic curves $\xi : [0, T] \rightarrow \mathcal{M}$ with $\xi(0) = p_0$ and the length of ξ is no greater than R .

We also need the following definition which was equivalently stated in (3.7) of [3], for $\mathcal{G}_2(p_0, R)$.

Definition 5. Let $\Omega \subseteq \mathcal{M}$ an open set and $R > 0$ a scalar constanst. A continuously differentiable $f : [0, R) \rightarrow \mathbb{R}$ is said to be a majorant function at a point $p_0 \in \Omega$ for a continuously differentiable vector field $X : \Omega \rightarrow T\mathcal{M}$ with respect to $\mathcal{G}_n(p_0, R)$ if $\nabla X(p_0)$ is nonsingular, $B(p_0, R) \subset \Omega$ and

$$\|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, a, b} - P_{\xi, a, 0} \nabla X(\xi(a))]\| \leq f'(\ell[\xi, 0, b]) - f'(\ell[\xi, 0, a]), \quad (5)$$

for all $\xi \in \mathcal{G}_n(p_0, R)$ with $a, b \in \text{dom}(\xi)$ and $0 \leq a \leq b$. Moreover, f satisfies the following conditions:

- h1.** $f(0) > 0$ and $f'(0) = -1$;
- h2.** f' is convex and strictly increasing;
- h3.** $f(t) = 0$ for some $t \in (0, R)$.

We also need of the following condition on the majorant condition f which will be considered to hold only when explicitly stated

- h4.** $f(t) < 0$ for some $t \in (0, R)$.

Remark 2. Since $f(0) > 0$ and f is continuous then condition **h4** implies condition **h3**.

The statement of our main result is:

Theorem 1. Let \mathcal{M} be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $\bar{\Omega}$ its closure, $X : \bar{\Omega} \rightarrow T\mathcal{M}$ a continuous vector field and continuously differentiable on Ω , $R > 0$ a scalar constant and $f : [0, R) \rightarrow \mathbb{R}$ a continuously differentiable function. Take $p_0 \in \Omega$. Suppose that $\nabla X(p_0)$ is nonsingular and f is a majorant function for X at p_0 with respect to $\mathcal{G}_3(p_0, R)$ satisfying **h4** and the inequality

$$\|\nabla X(p_0)^{-1}X(p_0)\| \leq f(0). \quad (6)$$

Define $\Gamma := \sup\{-f(t) : t \in [0, R)\}$. Let $0 \leq \rho < \Gamma/2$ and $g : [0, R - \rho) \rightarrow \mathbb{R}$,

$$g(t) := \frac{1}{|f'(\rho)|}[f(t + \rho) + 2\rho]. \quad (7)$$

Then g has a smallest zero $t_{*,\rho} \in (0, R - \rho)$, the sequences generated by Newton's Method for solving the equation $X(p) = 0$ and the equation $g(t) = 0$, with starting point q_0 , for any $q_0 \in B[p_0, \rho]$, and $t_0 = 0$, respectively,

$$q_{k+1} = \exp_{q_k}(-\nabla X(q_k)^{-1}X(q_k)), \quad t_{k+1} = t_k - \frac{g(t_k)}{g'(t_k)}, \quad k = 0, 1, \dots \quad (8)$$

are well defined, $\{q_k\}$ is contained in $B(q_0, t_{*,\rho})$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_{*,\rho})$ and converges to $t_{*,\rho}$. Moreover, $\{q_k\}$ and $\{t_k\}$ satisfy the inequalities

$$d(q_k, q_{k+1}) \leq t_{k+1} - t_k, \quad k = 0, 1, \dots, \quad (9)$$

$$d(q_k, q_{k+1}) \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} d(q_{k-1}, q_k)^2 \leq \frac{D^-g'(t_{*,\rho})}{-2g'(t_{*,\rho})} d(q_{k-1}, q_k)^2, \quad k = 1, 2, \dots \quad (10)$$

and $\{q_k\}$ converges to $p_* \in B[q_0, t_{*,\rho}]$ such that $X(p_*) = 0$. Furthermore, $\{q_k\}$ and $\{t_k\}$ satisfy the inequalities

$$d(q_k, p_*) \leq t_{*,\rho} - t_k, \quad t_{*,\rho} - t_{k+1} \leq \frac{1}{2}(t_{*,\rho} - t_k), \quad k = 0, 1, \dots, \quad (11)$$

the convergence of $\{q_k\}$ and $\{t_k\}$ to p_* and $t_{*,\rho}$, respectively, are Q -quadratic as follow

$$\limsup_{k \rightarrow \infty} \frac{d(p_{k+1}, p_*)}{d(q_k, p_*)^2} \leq \frac{D^-g'(t_{*,\rho})}{-2g'(t_{*,\rho})}, \quad t_{*,\rho} - t_{k+1} \leq \frac{D^-g'(t_{*,\rho})}{-2g'(t_{*,\rho})} (t_{*,\rho} - t_k)^2, \quad k = 0, 1, \dots \quad (12)$$

and p_* is the unique singularity of X in $B(p_0, \bar{\tau})$, where $\bar{\tau} \geq t_*$ is defined as

$$\bar{\tau} := \sup\{t \in [t_*, R) : f(t) \leq 0\}.$$

To prove the above theorem we need some previous results. First, in the next section, we prove a particular instance of this theorem, and then, in the Section 3.5 we prove Theorem 1.

3 Kantorovich's Theorem on Newton's Method

In this section we will prove an affine invariant version of Kantorovich's Theorem on Newton's Method, it is a particular instance of Theorem 1, namely, the case $\rho = 0$. We will use this theorem for proving Theorem 1. The main results of this section are the bounds, depending on the majorant function, for the Q -quadratic convergence of the Newton's Method, which gives an additional contribution for improving the results of Alvarez et al. in [3], Ferreira and Svaiter in [15] and Li and Wang in [25].

Theorem 2. Let \mathcal{M} be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $\bar{\Omega}$ its closure, $X : \bar{\Omega} \rightarrow T\mathcal{M}$ a continuous vector field and continuously differentiable on Ω , $R > 0$ a scalar constant and $f : [0, R) \rightarrow \mathbb{R}$ a continuously differentiable function. Take $p_0 \in \Omega$. Suppose that $\nabla X(p_0)$ is nonsingular and f is a majorant function for X at p_0 with respect to $\mathcal{G}_2(p_0, R)$ satisfying the inequality

$$\|\nabla X(p_0)^{-1}X(p_0)\| \leq f(0). \quad (13)$$

Then f has a smallest zero $t_* \in (0, R)$, the sequences generated by Newton's Method for solving the equations $X(p) = 0$ and $f(t) = 0$, with starting point p_0 and $t_0 = 0$, respectively,

$$p_{k+1} = \exp_{p_k}(-\nabla X(p_k)^{-1}X(p_k)), \quad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \quad k = 0, 1, \dots \quad (14)$$

are well defined, $\{p_k\}$ is contained in $B(p_0, t_*)$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$ and converge to t_* and satisfy the inequalities

$$d(p_{k+1}, p_k) \leq t_{k+1} - t_k, \quad d(p_{k+1}, p_k) \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} d(p_k, p_{k-1})^2, \quad (15)$$

for all $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively. Moreover, $\{p_k\}$ converge to $p_* \in B[p_0, t_*]$ such that $X(p_*) = 0$,

$$d(p_*, p_k) \leq t_* - t_k, \quad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \dots \quad (16)$$

and, therefore, $\{t_k\}$ converges Q -linearly to t_* and $\{p_k\}$ converge R -linearly to p_* . If, additionally, f satisfies **h4** then the following inequalities hold:

$$d(p_{k+1}, p_k) \leq \frac{D^- f'(t_*)}{-2f'(t_*)} d(p_k, p_{k-1})^2, \quad t_{k+1} - t_k \leq \frac{D^- f'(t_*)}{-2f'(t_*)} (t_k - t_{k-1})^2, \quad k = 1, 2, \dots, \quad (17)$$

and, as a consequence, $\{p_k\}$ and $\{t_k\}$ converge Q -quadratically to p_* and t_* , respectively, as follow

$$\limsup_{k \rightarrow \infty} \frac{d(p_*, p_{k+1})}{d(p_*, p_k)^2} \leq \frac{D^- f'(t_*)}{-2f'(t_*)}, \quad t_* - t_{k+1} \leq \frac{D^- f'(t_*)}{-2f'(t_*)} (t_* - t_k)^2, \quad k = 0, 1, \dots, \quad (18)$$

and p_* is the unique singularity of X in $B(p_0, \bar{\tau})$, where $\bar{\tau} \geq t_*$ is defined as

$$\bar{\tau} := \sup\{t \in [t_*, R) : f(t) \leq 0\}.$$

Henceforward we assume that all assumptions in above theorem hold. In this section, we will prove all the statements in Theorem 2 regarding to the majorant function and the real sequence $\{t_k\}$ associated. The main relationships between the majorant function and the vector field will be also established.

3.1 The majorant function

In this subsection we will study the majorant function f and prove all results regarding only the real sequence $\{t_k\}$ defined by Newton's method applied to the majorant function f . Define

$$\bar{t} := \sup\{t \in [0, R) : f'(t) < 0\}. \quad (19)$$

Proposition 2. *The majorant function f has a smallest root $t_* \in (0, R)$, is strictly convex and*

$$f(t) > 0, \quad f'(t) < 0, \quad t < t - f(t)/f'(t) < t_*, \quad \forall t \in [0, t_*]. \quad (20)$$

Moreover, $f'(t_*) \leq 0$ and

$$f'(t_*) < 0 \iff \exists t \in (t_*, R); f(t) < 0. \quad (21)$$

If, additionally, f satisfies condition **h4** then the following statements hold:

i) $f'(t) < 0$ for any $t \in [0, \bar{t}]$;

ii) $0 < t_* < \bar{t} \leq R$;

iii) $0 < \Gamma < \bar{t}$, where $\Gamma := -\lim_{t \rightarrow \bar{t}-} f(t)$.

iv) If $0 \leq \rho < \Gamma/2$ then $\rho < \bar{t}/2 < \bar{t}$ and $f'(\rho) < 0$.

Proof. See Propositions 2.3 and 5.2 of [17] and Proposition 3 of [15]. \square

In view of the second inequality in (20), Newton iteration is well defined in $[0, t_*)$. Let us call it $n_f : [0, t_*) \rightarrow \mathbb{R}$,

$$n_f(t) := t - f(t)/f'(t). \quad (22)$$

Proposition 3. *Newton iteration n_f maps $[0, t_*)$ into $[0, t_*)$ and there hold:*

$$t < n_f(t), \quad t_* - n_f(t) \leq \frac{1}{2}(t_* - t), \quad \forall t \in [0, t_*]. \quad (23)$$

If f also satisfies **(h4)**, i.e., $f'(t_*) < 0$, then

$$t_* - n_f(t) \leq \frac{D^- f'(t_*)}{-2f'(t_*)}(t_* - t)^2, \quad \forall t \in [0, t_*]. \quad (24)$$

Proof. See Proposition 4 of [16]. \square

The next two results follow from above proposition.

Corollary 1. *Take any $\tau_0 \in [0, t_*)$ and define, inductively, $\tau_{k+1} = n_f(\tau_k)$, $k = 0, 1, \dots$. The sequence $\{\tau_k\}$ is well defined, is strictly increasing, is contained in $[0, t_*)$ and converges Q -linearly to t_* as follows*

$$t_* - \tau_{k+1} \leq \frac{1}{2}(t_* - \tau_k), \quad k = 0, 1, \dots$$

In particular, the definition (14) of $\{t_k\}$ in Theorem 2 is equivalent to the following one

$$t_0 = 0, \quad t_{k+1} = n_f(t_k), \quad k = 0, 1, \dots \quad (25)$$

and there holds

Corollary 2. *The sequence $\{t_k\}$ is well defined, is strictly increasing, is contained in $[0, t_*)$ and converges Q -linearly to t_* as follows*

$$t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \dots$$

If f also satisfies **h4**, then the following inequality holds

$$t_{k+1} - t_k \leq \frac{D^- f'(t_*)}{-2f'(t_*)}(t_k - t_{k-1})^2, \quad k = 1, 2, \dots, \quad (26)$$

and, as a consequence, $\{t_k\}$ converges Q -quadratically to t_* as follow

$$t_* - t_{k+1} \leq \frac{D^- f'(t_*)}{-2f'(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \dots, \quad (27)$$

3.2 Relationship between the majorant function and the vector field

In this subsection we will establish the main relationship between the majorant function and the vector field necessities to prove Theorem 2.

Proposition 4. *Let $\xi \in \mathcal{G}_2(p_0, R)$. If $\ell[\xi, 0, s] \leq t < \bar{t}$ then $\nabla X(\xi(s))$ is nonsingular and the following inequality holds*

$$\|\nabla X(\xi(s))^{-1} P_{\xi,0,s} \nabla X(p_0)\| \leq \frac{1}{|f'(\ell[\xi, 0, s])|} \leq \frac{1}{|f'(t)|}.$$

Proof. Using Definition 5 and Lemma 3, the proof follows the same pattern of Proposition 3.4 of [16], see also Lemma 4.2. of [3]. \square

Newton iteration at a point happens to be a zero of the linearization at such a point. Therefore, we study the linearization error of the vector field and the associated majorant function. The formal definitions of these errors are:

Definition 6. *Let $f : [0, R) \rightarrow \mathbb{R}$ be a continuously differentiable function. The linearization error of f is defined by*

$$e(a, b) := f(b) - [f(t) + f'(a)(b - a)], \quad \forall a, b \in [0, R). \quad (28)$$

Definition 7. *Let \mathcal{M} be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set, $X : \Omega \rightarrow T\mathcal{M}$ a continuously differentiable vector field and $a, b \in [0, R)$. The linearization error of X on a geodesic $\zeta : [a, b] \rightarrow \Omega$ is defined by*

$$E(\zeta(a), \zeta(b)) := X(\zeta(b)) - P_{\zeta,a,b} [X(\zeta(a)) + (b - a) \nabla X(\zeta(a)) \zeta'(a)]. \quad (29)$$

In the next result we compare linearization error of the vector field with the linearization error of the majorant function associated.

Lemma 4. *Let $\xi \in \mathcal{G}_2(p_0, R)$ be a curve passing through $p = \xi(a)$ and $q = \xi(b)$ such that $\xi|_{[a,b]}$ is a geodesic and $0 \leq a \leq b$. Take $0 \leq t < x < R$. If $\ell[\xi, 0, a] \leq t$ and $\ell[\xi, a, b] \leq x - t$, then*

$$\|\nabla X(p_0)^{-1} P_{\xi,b,0} E(p, q)\| \leq e(t, x) \frac{\ell[\xi, a, b]^2}{(x - t)^2}.$$

As a consequence, the following inequality holds: $\|\nabla X(p_0)^{-1} P_{\xi,b,0} E(p, q)\| \leq e(t, x)$.

Proof. Definition 7 with $\zeta = \xi|_{[a,b]}$ and properties of parallel transport in (2) imply

$$E(p, q) = P_{\xi,a,b} [P_{\xi,b,a} X(q) - X(p) - (b - a) \nabla X(p) \xi'(a)].$$

Hence, using Lemma 1 and that $\xi'(s) = P_{\xi,a,s} \xi'(a)$, the last equality becomes

$$E(p, q) = P_{\xi,a,b} \int_a^b [P_{\xi,s,a} \nabla X(\xi(s)) P_{\xi,a,s} - \nabla X(p)] \xi'(a) ds,$$

which is equivalent to

$$\nabla X(p_0)^{-1} P_{\xi,b,0} E(p, q) = \int_a^b \nabla X(p_0)^{-1} [P_{\xi,s,0} \nabla X(\xi(s)) P_{\xi,a,s} - P_{\xi,a,0} \nabla X(p)] \xi'(a) ds.$$

Since $\xi : [a, b] \rightarrow \mathcal{M}$ is a geodesic joining p and q we have $\|\xi'(a)\| = \ell[\xi, a, b]/(b - a)$. Thus last equality implies

$$\begin{aligned} \|\nabla X(p_0)^{-1} P_{\xi, b, 0} E(p, q)\| &\leq \\ &\int_a^b \|\nabla X(p_0)^{-1} [P_{\xi, s, 0} \nabla X(\xi(s)) P_{\xi, a, s} - P_{\xi, a, 0} \nabla X(p)]\| \frac{\ell[\xi, a, b]}{b - a} ds. \end{aligned} \quad (30)$$

Because $a \leq s \leq b$, using the assumptions $\ell[\xi, 0, a] < t$ and $\ell[\xi, a, b] \leq x - t$ we have

$$\ell[\xi, 0, s] \leq \ell[\xi, 0, a] + \ell[\xi, a, b] \leq x < R,$$

and as $\xi : [0, s] \rightarrow \mathcal{M}$ is a piecewise geodesic curves joining the points p_0 to $\xi(s)$ through p , i. e., $\xi \in \mathcal{G}_2(p_0, R)$, we may use the majorant condition in Definition 5 with $b = s$ and $q = \xi(s)$ together with inequality in (30) to conclude that

$$\|\nabla X(p_0)^{-1} P_{\xi, b, 0} E(p, q)\| \leq \int_a^b [f'(\ell[\xi, 0, s]) - f'(\ell[\xi, 0, a])] \frac{\ell[\xi, a, b]}{b - a} ds.$$

Using convexity of f' , $\ell[\xi, 0, a] \leq t$, $\ell[\xi, a, b] \leq x - t$, $x < R$ and Proposition 1 we have

$$\begin{aligned} f'(\ell[\xi, 0, s]) - f'(\ell[\xi, 0, a]) &= f'(\ell[\xi, 0, a] + \ell[\xi, a, s]) - f'(\ell[\xi, 0, a]) \\ &\leq f'(t + \ell[\xi, a, s]) - f'(t) \\ &= f'\left(t + \frac{s - a}{b - a} \ell[\xi, a, b]\right) - f'(t) \\ &\leq \left[f'\left(t + \frac{s - a}{b - a} (x - t)\right) - f'(t)\right] \frac{\ell[\xi, a, b]}{x - t}. \end{aligned}$$

Therefore, combining two last inequality we obtain that

$$\|\nabla X(p_0)^{-1} P_{\xi, b, 0} E(p, q)\| \leq \int_a^b \left[f'\left(t + \frac{s - a}{b - a} (x - t)\right) - f'(t)\right] \frac{\ell[\xi, a, b]^2}{(x - t)(b - a)} ds.$$

After performing the integral and some algebraic manipulations the above inequality becomes

$$\|\nabla X(p_0)^{-1} P_{\xi, b, 0} E(p, q)\| \leq [f(x) - f(t) - f'(t)(x - t)] \frac{\ell[\xi, a, b]^2}{(x - t)^2},$$

which, Definition 6, implies the desired inequality. \square

Proposition 4 guarantees, in particular, that $\nabla X(p)$ is nonsingular at $p \in B(p_0, t_*)$ and, consequently, the *Newton's iteration* is well defined in $B(p_0, t_*)$. Let us call it $N_X : B(p_0, t_*) \rightarrow \mathcal{M}$,

$$N_X(p) := \exp_p(-\nabla X(p)^{-1} X(p)). \quad (31)$$

One can apply a *single* Newton's iteration on any $p \in B(p_0, t_*)$ to obtain the point $N_X(p)$ which may not be contained to $B(p_0, t_*)$, or even may not be in the domain of X . Hence, this is enough to guarantee the well-definedness of only one iteration. To ensure that Newtonian iteration may be repeated indefinitely, we need some additional definitions and results. First, we define some subsets of $B(p_0, t_*)$ in which, as we shall prove, Newton iteration (31) is “well behaved”:

$$K(t) := \left\{ p \in \Omega : d(p_0, p) \leq t, \quad \|\nabla X(p)^{-1} X(p)\| \leq -\frac{f(t)}{f'(t)} \right\}, \quad t \in [0, t_*]. \quad (32)$$

$$K := \bigcup_{t \in [0, t_*]} K(t), \quad (33)$$

In (32), $0 \leq t < t_* \leq \bar{t}$, hence using Proposition 2 and Proposition 4 we conclude that $f'(t) \neq 0$ and $\nabla X(p)$ is nonsingular in $B[p_0, t] \subset B[p_0, t_*)$, respectively. Therefore the above definitions are consistent. It is worth point out that the above sets appeared for the first time in [15]; see also [16].

Lemma 5. *For each $t \in [0, t_*)$ and each $p \in K(t)$ there hold:*

- i) $\|\nabla X(p)^{-1}X(p)\| \leq -\frac{f(t)}{f'(t)}$;
- ii) $d(p_0, p) + \|\nabla X(p)^{-1}X(p)\| \leq n_f(t) < t_*$. As a consequence, $d(p_0, N_X(p)) \leq n_f(t) < t_*$.
- iii) $\|\nabla X(N_X(p))^{-1}X(N_X(p))\| \leq -\frac{f(n_f(t))}{f'(n_f(t))} \left[\frac{\|\nabla X(p)^{-1}X(p)\|}{-f(t)/f'(t)} \right]^2$.

Proof. Let $t \in [0, t_*)$, $p \in K(t)$. Using definition of the set $K(t)$ in (32) the item **i** follows.

Using Proposition 3 and definition of $K(t)$ in (32) to obtain that $d(p_0, p) \leq t$ and $n_f(t) < t_*$, respectively. Hence, the proof of the first part of item **ii** follows by combination of two last inequalities with item **i** and definition of n_f in (22). For proving the second part of item **ii** use triangular inequality to obtain $d(p_0, N_X(p)) \leq d(p_0, p) + d(p, N_X(p))$, definition in (31) and then first part.

We are going to prove item **iii**. Let $\xi : [0, 2] \rightarrow \mathcal{M}$ a piecewise geodesic curve obtained by concatenation of a minimizing geodesic $\xi|_{[0,1]}$ joining p_0 and p and the geodesic curve $\xi|_{[1,2]}$ defined by

$$\xi(t) = \exp_p((1-t)\nabla X(p)^{-1}X(p)). \quad (34)$$

Note that $\xi \in \mathcal{G}_2(p_0, R)$. From definition of the piecewise geodesic curve ξ and definitions in (31) and (34) we have

$$\ell[\xi, 0, 2] = d(p_0, p) + \|\nabla X(p)^{-1}X(p)\|.$$

Since $\xi(2) = N_X(p)$, using last equality, first inequality in item **ii** and Proposition 4, by taking into account that the derivative f' is increasing and negative in $[0, \bar{t})$, we conclude that $\nabla X(N_X(p))$ is nonsingular and there holds

$$\|\nabla X(N_X(p))^{-1}P_{\xi,0,2}\nabla X(p_0)\| \leq \frac{1}{|f'(d(p_0, p) + \|\nabla X(p)^{-1}X(p)\|)|} \leq \frac{1}{|f'(n_f(t))|}. \quad (35)$$

On the other hand, as $\ell[\xi, 1, 2] = \|\nabla X(p)^{-1}X(p)\|$, combining item **i** with definition of n_f in (22) we obtain $\ell[\xi, 1, 2] \leq n_f(t) - t$. Since second part in item **ii** implies $d(p_0, N_X(p)) \leq n_f(t) < t_*$. Thus, we may apply Lemma 4 with $x = n_f(t)$ and $q = N_X(p)$ to conclude that

$$\|\nabla X(p_0)^{-1}P_{\xi,2,0}E(p, N_X(p))\| \leq e(t, n_f(t)) \frac{\|\nabla X(p)^{-1}X(p)\|^2}{(n_f(t) - t)^2}. \quad (36)$$

We know that $N_X(p)$ belongs to the domain of X . Hence, Newton's iterations in (31), linearization error in Definition 7 with $\zeta = \xi|_{[1,2]}$ and (34) yield

$$E(p, N_X(p)) = X(N_X(p)) - P_{\xi,1,2} [X(p) + \nabla X(p) (-\nabla X(p)^{-1}X(p))],$$

which is equivalent to $E(p, N_X(p)) = X(N_X(p))$. Thus, using this equality we obtain after simples algebraic manipulation that

$$\nabla X(N_X(p))^{-1}X(N_X(p)) = \nabla X(N_X(p))^{-1}P_{\xi,0,2}\nabla X(p_0)\nabla X(p_0)^{-1}P_{\xi,2,0}E(p, N_X(p)).$$

Taking norm in last equality and using the inequalities (35) and (36) we easily conclude that

$$\|\nabla X(N_X(p))^{-1}X(N_X(p))\| \leq \frac{e(t, n_f(t))}{|f'(n_f(t))|} \frac{\|\nabla X(p)^{-1}X(p)\|^2}{(n_f(t) - t)^2}.$$

Finally, since $n_f(t)$ belongs to the domain of f , using the definitions of Newton iterations on (22) and definition of the linearization error in (28), we obtain $f(n_f(t)) = e(t, n_f(t))$ which combined with $n_f(t) - t = f(t)/f'(t)$ and last inequality implies the desired result. Therefore, the proof of the lemma is concluded. \square

Lemma 6. *For each $t \in [0, t_*)$ the following inclusions hold: $K(t) \subset B(p_0, t_*)$ and*

$$N_X(K(t)) \subset K(n_f(t)).$$

As a consequence, $K \subset B(p_0, t_)$ and $N_X(K) \subset K$.*

Proof. The first inclusion follows trivially from the definition of $K(t)$ in (32). Combining items **i** and **iii** of Lemma 5 we have

$$\|\nabla X(N_X(p))^{-1}X(N_X(p))\| \leq \frac{f(n_f(t))}{|f'(n_f(t))|}.$$

Therefore, the second inclusion of the lemma follows from combination of last inequality in item **ii** of Lemma 5, last inequality and definition of $K(t)$. The first inclusion on the second sentence follows trivially from definitions (32) and (33). To verify the last inclusion, take $p \in K$. Then $p \in K(t)$ for some $t \in [0, t_*)$. Using the first part of the lemma, we conclude that $N_X(p) \subseteq K(n_f(t))$. To end the proof, note that $n_f(t) \in [0, t_*)$ and use the definition of K in (33). \square

We end this session limiting the derivative of the vector field by the derivative of the majorant function.

Proposition 5. *If $d(p_0, p) \leq t < R$ then $\|\nabla X(p)\| \leq \|\nabla X(p_0)\|(2 + f'(t))$.*

Proof. Let $\xi : [0, 1] \rightarrow \mathcal{M}$ is a minimizing geodesic joining p_0 to p . After some algebraic manipulations we have

$$\begin{aligned} \|\nabla X(p_0)^{-1}P_{\xi,1,0}\nabla X(p)\| &= \|\nabla X(p_0)^{-1}[P_{\xi,1,0}\nabla X(p)P_{\xi,0,1} - \nabla X(p_0) + \nabla X(p_0)]\| \\ &\leq \|\nabla X(p_0)^{-1}[P_{\xi,1,0}\nabla X(p)P_{\xi,0,1} - \nabla X(p_0)]\| + \|I_{p_0}\|. \end{aligned}$$

Since ξ is a minimizing geodesic joining p_0 to p we have $\ell[\xi, 0, 1] = d(p_0, p)$. Thus, using that f is a majorant function at a point p_0 for the vector field X , above inequality yields

$$\|\nabla X(p_0)^{-1}P_{\xi,1,0}\nabla X(p)\| \leq f'(d(p_0, p)) - f'(0) + 1 \leq 2 + f'(t),$$

because $d(p_0, p) \leq t$ and f' is an increasing function. Finally, using last inequality and taking into account that

$$\|\nabla X(p)\| \leq \|\nabla X(p_0)\| \|\nabla X(p_0)^{-1}P_{\xi,1,0}\nabla X(p)\|,$$

the desired inequality follows. \square

3.3 Convergence

In this section we establish all the convergence results stated in Theorem 2 related to $\{p_k\}$, the sequence generated by Newton's Method, namely, the convergence of $\{p_k\}$ to a zero of X , the bounds in (15), (16), (17) and (18). For establish these results we will combine conveniently the results of the previous section. We begin with the following result:

Proposition 6. *Let $\{z_k\}$ be a sequence in \mathcal{M} and $C > 0$. If $\{z_k\}$ converges to z_* and satisfies*

$$d(z_k, z_{k+1}) \leq C d(z_{k-1}, z_k)^2, \quad k = 1, 2, \dots \quad (37)$$

then $\{z_k\}$ converges Q -quadratically to z_ as follows*

$$\limsup_{k \rightarrow \infty} \frac{d(z_{k+1}, z_*)}{d(z_k, z_*)^2} \leq C.$$

Proof. The proof follows the same pattern as the proof of Proposition 1.2 of [13]. \square

Using equality in (14) and (31), the sequence $\{p_k\}$ generated by Newton's Method satisfies

$$p_{k+1} = N_X(p_k), \quad k = 0, 1, \dots \quad (38)$$

This equivalent definition of the Newton's sequence $\{p_k\}$ allow us to use the results of the previous section to establishes its properties of convergence.

Corollary 3. *The sequence $\{p_k\}$ is well defined, is contained in $B(p_0, t_*)$ and satisfies the inequalities in (15). Moreover, $\{p_k\}$ converges to a point $p_* \in B[p_0, t_*]$ satisfying $X(p_*) = 0$ and its convergence rate is R -linear as in (16). If, additionally, f satisfies **h4** then the inequality (17) holds and, consequently, $\{p_k\}$ converges Q -quadratically to p_* as in (18).*

Proof. We are going to prove that the sequence $\{p_k\}$ is well defined. First note that, combining (32), (13) and **h1** we have

$$p_0 \in K(0) \subset K, \quad (39)$$

where the second inclusion follows trivially from (33). Using the above inclusion, the inclusion $N_X(K) \subset K$ in Lemma 6 and (38) we conclude that $\{p_k\}$ is well defined and rests in K . From the first inclusion on second part of the Lemma 6 we have trivially that $\{p_k\}$ is contained in $B(p_0, t_*)$.

Now we are going to prove the inequalities in (15). First we will prove, by induction that

$$p_k \in K(t_k), \quad k = 0, 1, \dots \quad (40)$$

The above inclusion for $k = 0$ follows from (39). Assume now that $p_k \in K(t_k)$. Thus, using Lemma 6, (38) and (25), we obtain that $p_{k+1} \in K(t_{k+1})$, which completes the induction proof of (40). Using definition of $\{t_k\}$ in (14), we have $-f(t_k)/f'(t_k) = t_{k+1} - t_k$. Hence combining definition of $\{p_k\}$ in (14) with (40) and item **i** of Lemma 5, we obtain

$$d(p_k, p_{k+1}) = \|\nabla X(p_k)^{-1} X(p_k)\| \leq t_{k+1} - t_k, \quad k = 0, 1, \dots \quad (41)$$

which is first inequality in (15). In order to prove the second inequality in (15), first note that $p_{k-1} \in K(t_{k-1})$, $p_k = N_X(p_{k-1})$ and $t_k = n_f(t_{k-1})$, for all $k = 0, 1, \dots$. Thus, apply item **iii** of Lemma 5 with $p = p_{k-1}$ and $t = t_{k-1}$ to obtain

$$d(p_k, p_{k+1}) \leq -\frac{f(t_k)}{f'(t_k)} \left[\frac{d(p_{k-1}, p_k)}{t_k - t_{k-1}} \right]^2,$$

which using second inequality in (14) yields the desired inequality.

To prove that $\{p_k\}$ converges to $p_* \in B[p_0, t_*]$ with $X(p_*) = 0$ and (15) holds, first note that as $\{t_k\}$ converges to t_* , the first inequality (15) implies

$$\sum_{k=k_0}^{\infty} d(p_{k+1}, p_k) \leq \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty, \quad (42)$$

for any $k_0 \in \mathbb{N}$. Hence, $\{p_k\}$ is a Cauchy sequence in $B(p_0, t_*)$ and, thus, converges to some $p_* \in B[p_0, t_*]$. Therefore, first inequality (15) also implies that $d(p_*, p_k) \leq t_* - t_k$ for any k . Hence, the inequality (15) holds and, as $\{t_k\}$ converges Q -linearly to t_* , $\{p_k\}$ converges R -linearly to p_* . For proving that $X(p_*) = 0$, note that first inequality in (15) implies that $d(p_0, p_k) \leq t_k - t_0 = t_k$. Thus using Proposition 5 we have

$$\|\nabla X(p_k)\| \leq \|\nabla X(p_0)\|(2 + f'(t_k)), \quad k = 0, 1, \dots,$$

which combining inclusion (40) and second inequality in (41) yields

$$\|X(p_k)\| \leq \|\nabla X(p_k)\| \|\nabla X(p_k)^{-1} X(p_k)\| \leq \|\nabla X(p_0)\|(2 + f'(t_k))(t_{k+1} - t_k), \quad k = 0, 1, \dots$$

Since X is continuous on $\bar{\Omega}$, $\{p_k\} \subset B(p_0, t_*) \subset \bar{\Omega}$, $\{p_k\}$ converges to $p_* \in \bar{\Omega}$, the result follows by taking limit as k goes to infinite in above inequality.

Now, we assume that **h4** holds. Thus, combining second inequality in (15) with (26), we obtain the inequality in (17). To establish the inequality in (18), use inequality in (17) and Proposition 6 with $z_k = p_k$ and $C = D^- f'(t_*)/(-2f'(t_*))$. Therefore, the proof is concluded. \square

3.4 Uniqueness

In this section we prove the last statement in Theorem 2, namely, the uniqueness of the singularity of the vector field in consideration. The results of this section generalize [16, Section 3.2] for a general majorant function, see also [3, Section 4.2].

Corollary 4. *Take $0 \leq t < t_*$ and $q \in K(t)$. Define*

$$\tau_0 = t, \quad \tau_{k+1} = \tau_k - f(\tau_k)/f'(\tau_k), \quad k = 0, 1, \dots$$

The sequence $\{q_k\}$ generated by Newton's method with starting point $q_0 = q$ is well defined and satisfies $q_k \in K(\tau_k)$, for all k . Furthermore, $\{\tau_k\}$ converges to t_ , $\{q_k\}$ converges to some $q_* \in B[p_0, t_*]$ a singular point of X and $d(q_k, q_*) \leq t_* - \tau_k$, for all k .*

Proof. The proof is a convenient combination of Lemma 6, Corollary 1 and Proposition 5, following the same pattern of Corollary 3.6 of [16]. \square

The next two lemmas are most important results we need to prove the uniqueness of solution. The idea of its proofs are similar to the corresponding results of [16], see also [3]. In this more general approach, some technical details related to the parallel transport and the majorant function (possibly non-quadratic) should be used.

Lemma 7. *Take $0 \leq t < t_*$ and $p \in K(t)$. Define for $\theta \in \mathbb{R}$*

$$\zeta(\theta) = \exp_p(-\theta \nabla X(p)^{-1} X(p)), \quad \tau(\theta) = t - \theta \frac{f(t)}{f'(t)}.$$

Then for $\theta \in [0, 1]$ we have $t \leq \tau(\theta) < t_$ and $\zeta(\theta) \in K(\tau(\theta))$.*

Proof. The proof follows the same pattern of [16, Lemma 3.7], see also [3, Lemma 4.4]. \square

Lemma 8. Take $0 \leq t < t_*$ and $p \in K(t)$. Suppose that $q_* \in B[p_0, t_*]$ is a singular point of X and $t + d(p, q_*) = t_*$. Then $d(p_0, p) = t$. Furthermore, $t < n_f(t) < t_*$, $N_X(p) \in K(n_f(t))$ and $n_f(t) + d(N_X(p), q_*) = t_*$.

Proof. The proof follows the same pattern of [16, Lemma 3.8], see also [3, Lemma 4.5]. \square

The proof of the next two results can be obtained by a simple adaptation of some arguments of [16, Corollary 3.9] and [16, Lemma 3.10], see also [3, Lemma 4.5] and [3, Section 4.2.2], we also omit their proofs.

Corollary 5. Suppose that $\tilde{q}_* \in B[p_0, t_*]$ is a singular point of X . If for some \tilde{t}, \tilde{q}

$$0 \leq \tilde{t} < t_*, \quad \tilde{q} \in K(\tilde{t}),$$

and $\tilde{t} + d(\tilde{q}, \tilde{q}_*) = t_*$, then $d(p_0, \tilde{q}_*) = t_*$.

Lemma 9. The sequence $\{p_k\}$ has limit p_* as the unique singular point of X in $B[p_0, t_*]$.

Lemma 10. Let $q \in B(p_0, R)$ and $\xi : [0, 1] \rightarrow \mathcal{M}$ a minimizing geodesic in $\mathcal{G}_1(p_0, R)$ joining p_0 to q . Then the following inequality holds:

$$-f(d(p_0, q)) \leq \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} X(q)\|.$$

As a consequence, p_* is the unique singularity of X in $B(p_0, \bar{\tau})$, where $\bar{\tau} := \sup\{t \in [t_*, R) : f(t) \leq 0\}$.

Proof. Applying second part of Lemma 4 with $p = p_0$, $a = 0$, $b = 1$, $t = 0$ and $x = d(p_0, q)$ we have

$$e(0, d(p_0, q)) \geq \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} E(p_0, q)\|.$$

From Definition 7, last inequality becomes

$$e(0, d(p_0, q)) \geq \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} X(q) - \nabla X(p_0)^{-1} X(p_0) - \xi'(0)\|.$$

Using triangular inequality in the right hand side of last inequality, it is easy to see that

$$e(0, d(p_0, q)) \geq \|\xi'(0)\| - \|\nabla X(p_0)^{-1} X(p_0)\| - \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} X(q)\|.$$

Combining Definition 6 with assumption (13) and taking into account that $\|\xi'(0)\| = d(p_0, q)$ and $f'(0) = -1$, we obtain from last inequality that

$$f(d(p_0, q)) - [f(0) + f'(0)d(p_0, q)] \geq d(p_0, q) - f(0) - \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} X(q)\|,$$

with is equivalent to the inequality of the lemma. Hence the first of the lemma is proved.

For the second part, first note that in the interval $(t_*, \bar{\tau})$ the sign of f is negative. Hence, first part of the lemma implies that there is no singularity of X in $B(p_0, \bar{\tau}) \setminus B[p_0, t_*]$. Therefore, from Lemma 9, the unique singularity of X in $B(p_0, \bar{\tau})$ is $p_* \in B[p_0, t_*]$. \square

3.5 Proof of Theorem 2

The proof of Theorem 2 follow by direct combination of Corollary 2, Corollary 3 with Lemma 10.

4 On the proof of the main theorem

In this section Theorem 3 will be used to prove a robust semi-local affine invariant theorem for Newton's method for finding a singularity of the vector field X , namely, Theorem 1. The following result will be needed.

Proposition 7. *Let $R > 0$ and $f : [0, R) \rightarrow \mathbb{R}$ a continuously differentiable function. Suppose that $p_0 \in \Omega$, f is a majorant function for X at p_0 with respect to $\mathcal{G}_3(p_0, R)$ and satisfies **h4**. If $0 \leq \rho < \Gamma/2$, where $\Gamma := \sup\{-f(t) : t \in [0, R)\}$, then for any $q_0 \in B[p_0, \rho]$ the derivative $\nabla X(q_0)$ is nonsingular. Moreover, the scalar function $g : [0, R - \rho) \rightarrow \mathbb{R}$,*

$$g(t) = \frac{1}{|f'(\rho)|} [f(t + \rho) + 2\rho],$$

*is a majorant function for X at q_0 with respect to $\mathcal{G}_2(q_0, R - \rho)$ and also satisfies condition **h4**.*

Proof. Since the domain of f is $[0, R)$ and $f'(\rho) < 0$ (see Proposition 2 item iv), we conclude that g is well defined. First we will prove that function g satisfies conditions **h1**, **h2**, **h3** and **h4**. Definition of g and $f'(\rho) < 0$ trivially imply $g'(0) = -1$. Since f is convex and $f'(0) = -1$ we have $f(t) + t \geq f(0) > 0$, for all $0 \leq t < R$, which, by using Proposition 2 item iv and that $0 \leq \rho$, yields $g(0) = [f(\rho) + 2\rho]/|f'(\rho)| > 0$, hence g satisfies **h1**. Using that f satisfies **h2**, we easily conclude that g also satisfies **h2**. Now, as $\rho < \Gamma/2$, using Proposition 2 item iii, we have

$$\lim_{t \rightarrow \bar{t} - \rho} g(t) = \frac{1}{|f'(\rho)|} (2\rho - \Gamma) < 0,$$

which implies that g satisfies **h4** and, as g is continuous and $g(0) > 0$, it also satisfies **h3**.

To complete the proof, it remains to prove that g satisfies (5). First of all, for any $q_0 \in B[p_0, \rho]$, from Proposition 2 item iv, we have $d(q_0, p_0) \leq \rho < \bar{t}$. Let $\eta : [0, 1] \rightarrow \mathcal{M}$ be the minimizing geodesic joining p_0 to q_0 . Since $\eta \in \mathcal{G}_1(p_0, R) \subset \mathcal{G}_2(p_0, R)$ and $d(p_0, q_0) = \ell[\eta, 0, 1] \leq \rho < \bar{t}$ we can apply Proposition 4 to obtain that $\nabla X(q_0)$ is nonsingular and

$$\|\nabla X(q_0)^{-1} P_{\eta, 0, 1} \nabla X(p_0)\| \leq \frac{1}{|f'(\rho)|}. \quad (43)$$

Because $B(p_0, R) \subseteq \Omega$, for any $q_0 \in B[p_0, \rho]$, we trivially have $B(q_0, R - \rho) \subset \Omega$. Let $\mu : [0, T] \rightarrow \mathcal{M}$ such that $\mu \in \mathcal{G}_2(q_0, R - \rho)$ and $c_0, c_1, c_2 \in [0, T]$ with $c_0 = 0 \leq c_1 \leq c_2 = T$ such that $\mu|_{[c_0, c_1]}$ is a minimizing geodesic and $\mu|_{[c_1, c_2]}$ is a geodesic. Take $a, b \in [0, T]$ with $0 \leq a \leq b$. Thus

$$\mu(a), \mu(b) \in B(q_0, R - \rho), \quad \ell[\mu, 0, a] + \ell[\mu, a, b] < R - \rho, \quad d(q_0, \mu(a)) = \ell[\mu, 0, a].$$

Using definitions of the curves η and μ , properties of the parallel transport, property of the norm and simple manipulation, we conclude that

$$\begin{aligned} & \|\nabla X(q_0)^{-1} [P_{\mu, b, 0} \nabla X(\mu(b)) P_{\mu, a, b} - P_{\mu, a, 0} \nabla X(\mu(a))]\| \leq \\ & \|\nabla X(q_0)^{-1} P_{\eta, 0, 1} \nabla X(p_0)\| \|\nabla X(p_0)^{-1} P_{\eta, 1, 0} [P_{\mu, b, 0} \nabla X(\mu(b)) P_{\mu, a, b} - P_{\mu, a, 0} \nabla X(\mu(a))]\|. \end{aligned} \quad (44)$$

Now we are going to estimate the second norm of the right hand side of above inequality. First, we define $\xi : [0, \hat{T}] \rightarrow \mathcal{M}$ a piecewise geodesic curve in $\mathcal{G}_3(p_0, R)$ as concatenation between the curves η and μ , i.e., take $\hat{c}_0 = 0 < \hat{c}_1 < \hat{c}_2 < \hat{c}_3 = \hat{T}$ such that

$$\xi|_{[\hat{c}_0, \hat{c}_1]} = \eta|_{[0, 1]}, \quad \xi|_{[\hat{c}_1, \hat{c}_2]} = \mu|_{[0, c_1]}, \quad \xi|_{[\hat{c}_2, \hat{c}_3]} = \mu|_{[c_1, c_2]}. \quad (45)$$

Definition of ξ in (45) and definition of curve μ imply that there exist $\hat{a}, \hat{b} \in \text{dom}(\xi)$ with $0 \leq \hat{a} \leq \hat{b}$ such that $\xi(\hat{a}) = \mu(a)$ and $\xi(\hat{b}) = \mu(b)$. Therefore, properties of parallel transport yield $P_{\eta,1,0}P_{\mu,b,0} = P_{\xi,\hat{b},0}$. Hence,

$$\begin{aligned} \left\| \nabla X(p_0)^{-1} P_{\eta,1,0} [P_{\mu,b,0} \nabla X(\mu(b)) P_{\mu,a,b} - P_{\mu,a,0} \nabla X(\mu(a))] \right\| = \\ \left\| \nabla X(p_0)^{-1} [P_{\xi,\hat{b},0} \nabla X(\xi(\hat{b})) P_{\xi,\hat{a},\hat{b}} - P_{\xi,\hat{a},0} \nabla X(\xi(\hat{a}))] \right\|. \end{aligned}$$

Since $\xi \in \mathcal{G}_3(p_0, R)$ and f is a majorant function for X at p_0 with respect to $\mathcal{G}_3(p_0, R)$, applying Definition 5 with $a = \hat{a}$ and $b = \hat{b}$, last equality becomes

$$\left\| \nabla X(p_0)^{-1} P_{\eta,1,0} [P_{\mu,b,0} \nabla X(\mu(b)) P_{\mu,a,b} - P_{\mu,a,0} \nabla X(\mu(a))] \right\| \leq f'(\ell[\xi, 0, \hat{b}]) - f'(\ell[\xi, 0, \hat{a}]). \quad (46)$$

Combining last inequality with (43), (44) and (46) we obtain

$$\begin{aligned} \left\| \nabla X(q_0)^{-1} [P_{\mu,b,0} \nabla X(\mu(b)) P_{\mu,a,b} - P_{\mu,a,0} \nabla X(\mu(a))] \right\| \leq \\ \frac{1}{|f'(\rho)|} \left[f'(\ell[\xi, 0, \hat{b}]) - f'(\ell[\xi, 0, \hat{a}]) \right]. \quad (47) \end{aligned}$$

Since f' is convex, the function $s \mapsto f'(t+s) - f'(s)$ is increasing for $t \geq 0$. Hence taking into account that definitions of ξ in (45) and μ imply $\ell[\xi, 0, \hat{a}] = \ell[\xi, 0, \hat{c}_1] + \ell[\xi, \hat{c}_1, \hat{a}] \leq \rho + \ell[\mu, 0, a]$ and $\ell[\xi, 0, \hat{b}] = \ell[\xi, 0, \hat{a}] + \ell[\xi, \hat{a}, \hat{b}] \leq \rho + \ell[\mu, 0, a] + \ell[\mu, a, b]$, we conclude that

$$f'(\ell[\xi, 0, \hat{b}]) - f'(\ell[\xi, 0, \hat{a}]) \leq f'(\rho + \ell[\mu, 0, a] + \ell[\mu, a, b]) - f'(\rho + \ell[\mu, 0, a]).$$

Since $\ell[\mu, 0, b] = \ell[\mu, 0, a] + \ell[\mu, a, b]$, combining inequality in (47) and last inequality with the definition of the function g we have

$$\left\| \nabla X(q_0)^{-1} [P_{\mu,b,0} \nabla X(\mu(b)) P_{\mu,a,b} - P_{\mu,a,0} \nabla X(\mu(a))] \right\| \leq g'(\ell[\mu, 0, b]) - g'(\ell[\mu, 0, a]),$$

implying that the function g satisfies (5), which complete the proof of the proposition. \square

Proposition 8. *Let $q \in B(p_0, R)$ and $\xi : [0, 1] \rightarrow \mathcal{M}$ a minimizing geodesic joining p_0 to q . Then the following inequality holds:*

$$\left\| \nabla X(p_0)^{-1} P_{\xi,1,0} X(q) \right\| \leq f(d(p_0, q)) + 2d(p_0, q). \quad (48)$$

Proof. Applying second part of Lemma 4 with $p = p_0$, $a = 0$, $b = 1$, $t = 0$ and $x = d(p_0, q)$ we have

$$e(0, d(p_0, q)) \geq \left\| \nabla X(p_0)^{-1} P_{\xi,1,0} E(p_0, q) \right\|.$$

From Definition 7 last inequality becomes

$$e(0, d(p_0, q)) \geq \left\| \nabla X(p_0)^{-1} P_{\xi,1,0} X(q) - \nabla X(p_0)^{-1} X(p_0) - \xi'(0) \right\|.$$

Using triangular inequality in the right hand side of last inequality, it is easy to see that

$$e(0, d(p_0, q)) \geq \left\| \nabla X(p_0)^{-1} P_{\xi,1,0} X(q) \right\| - \left\| \nabla X(p_0)^{-1} X(p_0) \right\| - \left\| \xi'(0) \right\|.$$

Combining Definition 6 with assumption (13) and taking into account that $\|\xi'(0)\| = d(p_0, q)$ and $f'(0) = -1$, we obtain from last inequality that

$$f(d(p_0, q)) - [f(0) + f'(0)d(p_0, q)] \geq \left\| \nabla X(p_0)^{-1} P_{\xi,1,0} X(q) \right\| - f(0) - d(p_0, q),$$

which is equivalent to the inequality of the lemma. Hence the lemma is proved. \square

4.1 Proof of Theorem 1

Proposition 7 claims that for any $q_0 \in B[p_0, \rho]$ the derivative $\nabla X(q_0)$ is nonsingular. Moreover, the scalar function $g : [0, R - \rho] \rightarrow \mathbb{R}$,

$$g(t) = \frac{1}{|f'(\rho)|} [f(t + \rho) + 2\rho], \quad (49)$$

is a majorant function for X at q_0 with respect to $\mathcal{G}_2(q_0, R - \rho)$ and also satisfies condition **h4**. Let $\xi : [0, 1] \rightarrow \mathcal{M}$ a minimizing geodesic joining p_0 to q_0 . Since item *iv* of Proposition 2 implies $\ell[\xi, 0, 1] = d(p_0, q_0) \leq \rho < \bar{t}$, thus Proposition 4 give us

$$\|\nabla X(q_0)^{-1} P_{\xi, 0, 1} \nabla X(p_0)\| \leq \frac{1}{|f'(\rho)|}.$$

Combining property of norm with last inequality and Proposition 8 with $q = q_0$, we have

$$\begin{aligned} \|\nabla X(q_0)^{-1} X(q_0)\| &\leq \|\nabla X(q_0)^{-1} P_{\xi, 0, 1} \nabla X(p_0)\| \|\nabla X(p_0)^{-1} P_{\xi, 1, 0} X(q_0)\| \\ &\leq \frac{1}{|f'(\rho)|} [f(d(q_0, p_0)) + 2d(q_0, p_0)]. \end{aligned}$$

As $f' \geq -1$, the function $t \mapsto f(t) + 2t$ is (strictly) increasing. Using this fact, above inequality, $d(p_0, q_0) \leq \rho$ and (49) we conclude that

$$\|\nabla X(q_0)^{-1} X(q_0)\| \leq g(0).$$

Therefore, last inequality allow us to apply Theorem 2 for X and the majorant function g at point q_0 for obtaining the desired result.

5 Special cases

Kantorovich's theorem under a majorant condition in Riemannian settings was used in [3], see also [25] to prove Kantorovich's theorem under Lipschitz condition in Riemannian manifolds [15], Smale's theorem [37] and Nesterov-Nemirovskii's theorem [31]. Using the ideas of [3] we present, as an application of Theorem 1, a robust version of these theorems.

5.1 Under Lipschitz's condition

Theorem 3. *Let \mathcal{M} be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $\bar{\Omega}$ its closure, $X : \bar{\Omega} \rightarrow T\mathcal{M}$ a continuous vector field and continuously differentiable on Ω . Take $p_0 \in \Omega$, $L > 0$, $\beta > 0$ and $R = \sup\{r > 0 : B(p_0, r) \subset \Omega\}$. Suppose that $\nabla X(p_0)$ is nonsingular, $B(p_0, 1/L) \subset \Omega$,*

$$\|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, a, b} - P_{\xi, a, 0} \nabla X(\xi(a))]\| \leq L \ell[\xi, a, b],$$

for all ξ in $\mathcal{G}_3(p_0, R)$ and $2\beta L < 1$. Moreover, assume that

$$\|\nabla X(p_0)^{-1} X(p_0)\| \leq \beta.$$

Let $0 \leq \rho < (1 - 2\beta L)/(4L)$ and $t_{*, \rho} = \left(1 - \rho L - \sqrt{1 - 2L(\beta + 2\rho)}\right)/L$. Then the sequence generated by Newton's Method for solving the equations $X(p) = 0$, with starting point q_0 , for any $q_0 \in B[p_0, \rho]$,

$$q_{k+1} = \exp_{q_k} (-\nabla X(q_k)^{-1} X(q_k)), \quad k = 0, 1, \dots$$

is well defined, $\{q_k\}$ is contained in $B(q_0, t_{*,\rho})$ and satisfy the inequality

$$d(q_k, q_{k+1}) \leq \frac{L}{2\sqrt{1-2L(\beta+2\rho)}} d(q_{k-1}, q_k)^2, \quad k = 1, 2, \dots$$

Moreover, $\{q_k\}$ converges to $p_* \in B[q_0, t_{*,\rho}]$ such that $X(p_*) = 0$ and the convergence is Q -quadratic as follows

$$\limsup_{k \rightarrow \infty} \frac{d(q_{k+1}, p_*)}{d(q_k, p_*)^2} \leq \frac{L}{2\sqrt{1-2L(\beta+2\rho)}}.$$

Furthermore, if $B(p_0, \tau) \subset \Omega$ then p_* is the unique singularity of X in $B(p_0, \tau)$, where $\tau := (1 + \sqrt{1-2\beta L})/L$.

Proof. The proof follows from Theorem 1 with the quadratic polynomial $f(t) = \frac{L}{2}t^2 - t + \beta$ as the majorant function to X with respect to $\mathcal{G}_3(p_0, 1/L)$ and $\Gamma = (1 - 2\beta L)/(4L)$. \square

5.2 Under Smale's condition

Theorem 4. Let \mathcal{M} be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \rightarrow T\mathcal{M}$ an analytic vector field. Let $p_0 \in \mathcal{M}$ be such that $\nabla X(p_0)$ is nonsingular and set $\beta := \|\nabla X(p_0)^{-1}X(p_0)\|$. Suppose

$$\alpha := \beta\gamma < 3 - 2\sqrt{2}, \quad \gamma := \sup_{n \geq 1} \left\| \frac{1}{n!} \nabla X(p_0)^{-1} \nabla^n X(p_0) \right\|^{1/(n-1)} < \infty,$$

$B(p_0, R) \subset \Omega$, where $R := (1 - 1/\sqrt{2})/\gamma$. Let $0 \leq \rho < [3 - 2\sqrt{2} - \alpha]/(2\gamma)$ and

$$t_{*,\rho} := \left(\alpha + 1 - 2\rho\gamma - \sqrt{(\alpha + 1 - 2\rho\gamma)^2 - 8\alpha - 8\rho\gamma(1 - \alpha)} \right) / (4\gamma).$$

Then the sequences generated by Newton's method for solving the equations $X(p) = 0$ with starting at q_0 , for any $q_0 \in B[p_0, \rho]$,

$$q_{k+1} = \exp_{q_k}(-\nabla X(q_k)^{-1}X(q_k)), \quad k = 0, 1, \dots$$

are well defined, $\{q_k\}$ is contained in $B[q_0, t_{*,\rho}]$ and satisfy the inequality

$$d(q_k, q_{k+1}) \leq \frac{\gamma}{(1 - \gamma(t_{*,\rho} + \rho))[2(1 - \gamma(t_{*,\rho} + \rho))^2 - 1]} d(q_{k-1}, q_k)^2, \quad k = 1, 2, \dots$$

Moreover, $\{q_k\}$ converges to $p_* \in B[p_0, t_{*,0}]$ such that $X(p_*) = 0$ and the convergence is Q -quadratic as follows

$$\limsup_{k \rightarrow \infty} \frac{d(q_{k+1}, p_*)}{d(q_k, p_*)^2} \leq \frac{\gamma}{(1 - \gamma(t_{*,\rho} + \rho))[2(1 - \gamma(t_{*,\rho} + \rho))^2 - 1]}.$$

Furthermore, p_* is the unique singularity of X in $B(p_0, R) \subset \Omega$.

We need the following results to prove the above theorem.

Lemma 11. Let \mathcal{M} be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \rightarrow T\mathcal{M}$ an analytic vector field. Suppose that $p_0 \in \Omega$, $\nabla X(p_0)$ is nonsingular and that $R \leq (1 - 1/\sqrt{2})\gamma^{-1}$. Then, for all $\zeta \in \mathcal{G}_3(p_0, R)$ there holds

$$\|\nabla X(p_0)^{-1}P_{\zeta,s,0}\nabla^2 X(\zeta(s))\| \leq (2\gamma)/(1 - \gamma\ell[\zeta, 0, s])^3.$$

Proof. The proof follows the same pattern of Lemma 5.3 of [3]. \square

Lemma 12. *Let \mathcal{M} be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \rightarrow T\mathcal{M}$ an analytic vector field. Suppose that $p_0 \in \Omega$ and $\nabla X(p_0)$ is nonsingular. If there exists an $f : [0, R) \rightarrow \mathbb{R}$ twice continuously differentiable such that*

$$\|\nabla X(p_0)^{-1} P_{\zeta, s, 0} \nabla^2 X(\zeta(s))\| \leq f''(\ell[\zeta, 0, s]), \quad (50)$$

for all $\zeta \in \mathcal{G}_3(p_0, R)$ and for all $s \in \text{dom}(\zeta)$, then X and f satisfy (5) with $n = 3$.

Proof. Let ζ be a curve of $\mathcal{G}_3(p_0, R)$, $a, b \in \text{dom}(\zeta)$ with $0 \leq a \leq b$. From Definition 4 there exist $c_0, c_1, c_2, c_3 \in [0, T]$ with $c_0 = 0 \leq c_1 \leq c_2 \leq c_3 = T$ such that $\xi_{[c_0, c_1]}$ and $\xi_{[c_1, c_2]}$ are minimizing geodesics and $\xi_{[c_2, c_3]}$ is a geodesic. We have six possibilities:

- $a, b \in [c_i, c_{i+1}]$ for $i = 0, 1, 2$;
- $a \in [c_i, c_{i+1}]$ and $b \in [c_{i+1}, c_{i+2}]$ for $i = 0, 1$;
- $a \in [c_0, c_1]$ and $b \in [c_2, c_3]$.

We are going to analyze the possibility $a \in [c_0, c_1]$ and $b \in [c_2, c_3]$, the others are similar. Since $a \in [c_0, c_1]$ and $\xi_{[c_0, c_1]}$ is geodesic, taking $v \in T_{\zeta(a)}\mathcal{M}$ and $Y \in \mathcal{X}(\mathcal{M})$ the vector field on ζ such that $\nabla_{\zeta'(s)} Y = 0$ and $Y(\zeta(a)) = v$, we may apply Lemma 2 to have

$$P_{\zeta, c_1, a} \nabla X(\zeta(c_1)) Y(\zeta(c_1)) = \nabla X(\zeta(a)) Y(\zeta(a)) + \int_a^{c_1} P_{\zeta, s, a} \nabla^2 X(\zeta(s)) (Y(\zeta(s)), \zeta'(s)) ds. \quad (51)$$

Using that $Y(\zeta(a)) = v$ and $Y(\zeta(c_1)) = P_{\zeta, a, c_1} v$, we obtain, after some algebraic manipulation in last equality, that

$$\begin{aligned} \nabla X(p_0)^{-1} [P_{\zeta, c_1, 0} \nabla X(\zeta(c_1)) P_{\zeta, a, c_1} - P_{\zeta, a, 0} \nabla X(\zeta(a))] v = \\ \int_a^{c_1} \nabla X(p_0)^{-1} P_{\zeta, s, 0} \nabla^2 X(\zeta(s)) (Y(\zeta(s)), \zeta'(s)) ds. \end{aligned}$$

Since $\|Y(\zeta(s))\| = \|v\|$ for all $s \in [a, c_1]$ and v is arbitrary, we conclude from Definition 3 that

$$\begin{aligned} \|\nabla X(p_0)^{-1} [P_{\zeta, c_1, 0} \nabla X(\zeta(c_1)) P_{\zeta, a, c_1} - P_{\zeta, a, 0} \nabla X(\zeta(a))]\| \leq \\ \int_a^{c_1} \|\nabla X(p_0)^{-1} P_{\zeta, s, 0} \nabla^2 X(\zeta(s))\| \|\zeta'(s)\| ds. \end{aligned}$$

Now, as $\|\zeta'(s)\| = \ell[\zeta, a, c_1]/(c_1 - a)$ and $\ell[\zeta, 0, s] = \ell[\zeta, 0, a] + ((c_1 - s)/(c_1 - a))\ell[\zeta, a, c_1] < R$ for all $s \in [a, c_1]$, using (50) we obtain, from the last inequality, that

$$\begin{aligned} \|\nabla X(p_0)^{-1} [P_{\zeta, c_1, 0} \nabla X(\zeta(c_1)) P_{\zeta, a, c_1} - P_{\zeta, a, 0} \nabla X(\zeta(a))]\| \leq \\ \int_a^{c_1} f'' \left(\ell[\zeta, 0, a] + \frac{c_1 - s}{c_1 - a} \ell[\zeta, a, c_1] \right) \frac{\ell[\zeta, a, c_1]}{c_1 - a} ds. \end{aligned}$$

Evaluating the latter integral, it follows that

$$\|\nabla X(p_0)^{-1} [P_{\xi, c_1, 0} \nabla X(\xi(c_1)) P_{\xi, a, c_1} - P_{\xi, a, 0} \nabla X(\xi(a))]\| \leq f'(\ell[\xi, 0, c_1]) - f'(\ell[\xi, 0, a]). \quad (52)$$

On the other hand, using that $\xi|_{[c_1, c_2]}$ is geodesic, similar arguments used above show that

$$\|\nabla X(p_0)^{-1} [P_{\xi, c_2, 0} \nabla X(\xi(c_2)) P_{\xi, c_1, c_2} - P_{\xi, c_1, 0} \nabla X(\xi(c_1))]\| \leq f'(\ell[\xi, 0, c_2]) - f'(\ell[\xi, 0, c_1]). \quad (53)$$

We may also use that $b \in [c_2, c_3]$ and $\xi|_{[c_2, c_3]}$ is geodesic to obtain the following inequality

$$\|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, c_2, b} - P_{\xi, c_2, 0} \nabla X(\xi(c_2))]\| \leq f'(\ell[\xi, 0, b]) - f'(\ell[\xi, 0, c_2]). \quad (54)$$

Now, taking into account that the parallel transport is an isometry, the triangular inequality yields

$$\begin{aligned} \|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, a, b} - P_{\xi, a, 0} \nabla X(\xi(a))]\| \leq \\ \|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, c_2, b} - P_{\xi, c_2, 0} \nabla X(\xi(c_2))]\| + \\ \|\nabla X(p_0)^{-1} [P_{\xi, c_2, 0} \nabla X(\xi(c_2)) P_{\xi, c_1, c_2} - P_{\xi, c_1, 0} \nabla X(\xi(c_1))]\| + \\ \|\nabla X(p_0)^{-1} [P_{\xi, c_1, 0} \nabla X(\xi(c_1)) P_{\xi, a, c_1} - P_{\xi, a, 0} \nabla X(\xi(a))]\|. \end{aligned}$$

Combining last inequality with (52), (53) and (54), it follows that

$$\|\nabla X(p_0)^{-1} [P_{\xi, b, 0} \nabla X(\xi(b)) P_{\xi, a, b} - P_{\xi, a, 0} \nabla X(\xi(a))]\| \leq f'(\ell[\xi, 0, b]) - f'(\ell[\xi, 0, a]),$$

which is the desired result. \square

Proof of Theorem 4. Since $\alpha < 3 - 2\sqrt{2}$, combining Lemma 11 and Lemma 12 we have that the analytic function $f : [0, R) \rightarrow \mathbb{R}$ defined by $f(t) = \beta - 2t + t/(1 - \gamma t)$ is a majorant function to X with respect to $\mathcal{G}_3(p_0, R)$. Hence, the proof follows from Theorem 1 with $\Gamma = (3 - 2\sqrt{2} - \alpha)/\gamma$. \square

5.3 Under Nesterov-Nemirovskii's condition

Theorem 5. Let $C \subset \mathbb{R}^n$ be a open convex set and $F : C \rightarrow \mathbb{R}$ be a strictly convex function, three times continuously differentiable. Take $x_0 \in C$ with $F''(x_0)$ nonsingular. Define the norm

$$\|u\|_{x_0} := \sqrt{\langle u, u \rangle_{x_0}}, \quad \forall u \in \mathbb{R}^n,$$

where $\langle u, v \rangle_{x_0} = a^{-1} \langle F''(x_0)u, v \rangle$ for all $u, v \in \mathbb{R}^n$ and some $a > 0$. Suppose that F is a -self-concordant, i.e., satisfies

$$|F'''(x)[h, h, h]| \leq 2a^{-1/2} (F''(x)[h, h])^{3/2}, \quad \forall x \in C, h \in \mathbb{R}^n,$$

$W_1(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < 1\} \subset C$ and there exists $\beta \geq 0$ such that

$$\|F''(x_0)^{-1} F'(x_0)\|_{x_0} \leq \beta < 3 - 2\sqrt{2}.$$

Let $0 \leq \rho < (3 - 2\sqrt{2} - \beta)/2$ and $t_{*, \rho} := \left(\alpha + 1 - 2\rho - \sqrt{(\alpha + 1 - 2\rho)^2 - 8\alpha - 8\rho(1 - \alpha)} \right) / 4$. Then the sequences generated by Newton's method for solving the equations $F'(x) = 0$ with starting at y_0 , for any $y_0 \in W_\rho[x_0] = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} \leq \rho\}$,

$$y_{k+1} = y_k - F''(y_k)^{-1} F'(y_k), \quad k = 0, 1, \dots$$

is well defined, $\{y_k\}$ is contained in $W_{t_{*, \rho}}[x_0] = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} \leq t_{*, \rho}\}$ and satisfy the inequality

$$\|y_{k+1} - y_k\| \leq \frac{1}{(1 - (t_{*, \rho} + \rho))[2(1 - (t_{*, \rho} + \rho))^2 - 1]} \|y_k - y_{k-1}\|^2, \quad k = 1, 2, \dots$$

Moreover, $\{y_k\}$ converges to $x_* \in W_{t_*,0}[x_0]$ such that $F'(x_*) = 0$ and the convergence is Q -quadratic as follows

$$\limsup_{k \rightarrow \infty} \frac{\|x_* - y_{k+1}\|}{\|x_* - y_k\|^2} \leq \frac{1}{(1 - (t_{*,\rho} + \rho))[2(1 - (t_{*,\rho} + \rho))^2 - 1]}.$$

Proof. Since $\alpha < 3 - 2\sqrt{2}$, combining Lemma 5.1 of [3] and Lemma 12 we have that the function $f : [0, R) \rightarrow \mathbb{R}$ defined by $f(t) = \beta - 2t + t/(1 - t)$ is a majorant function to F' with respect to $\mathcal{G}_3(x_0, R)$. Hence, the proof follows from Theorem 1 with $\Gamma = 3 - 2\sqrt{2} - b$. \square

6 Final remark

Let us present some computational aspects of Newton's method in Riemannian settings for solving the equation (4). Note that the first equality in (8) is equivalent to

$$q_{k+1} = \exp_{q_k} S_k, \quad \nabla X(q_k)^{-1} S_k = -X(q_k), \quad k = 0, 1, \dots \quad (55)$$

Since the solution of the linear systems in (55) for large systems is computationally expensive, namely, at each iteration the derivative at q_k must be computed and stored. Besides, the solution of the linear system in (55) is required. To circumvent these drawbacks, we propose the inexact Newton's method: given an initial point q_0 , the method generates a sequence $\{q_k\}$ as follows:

$$q_{k+1} = \exp_{q_k} S_k, \quad \nabla X(q_k)^{-1} S_k = -X(q_k) + r_k, \quad \|r_k\| \leq \theta_k \|X(q_k)\| \quad k = 0, 1, \dots$$

for a suitable forcing sequence $\{\theta_k\}$, which is used to control the level of accuracy. Therefore, solutions of practical problems are obtained by computational implementations of the inexact Newton-like methods. The analysis of these methods under majorant condition will be done in the near future.

References

- [1] P.-A. Absil, L. Amodei, and G. Meyer. Two Newton methods on the manifold of fixed-rank matrices endowed with Riemannian quotient geometries. *Comput. Statist.*, 29(3-4):569–590, 2014.
- [2] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.*, 22(3):359–390, 2002.
- [3] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.
- [4] S. Amat, S. Busquier, R. Castro, and S. Plaza. Third-order methods on Riemannian manifolds under Kantorovich conditions. *J. Comput. Appl. Math.*, 255:106–121, 2014.
- [5] I. K. Argyros. An improved unifying convergence analysis of Newton's method in Riemannian manifolds. *J. Appl. Math. Comput.*, 25(1-2):345–351, 2007.
- [6] I. K. Argyros and S. Hilout. Newton's method for approximating zeros of vector fields on Riemannian manifolds. *J. Appl. Math. Comput.*, 29(1-2):417–427, 2009.

- [7] I. K. Argyros and Á. A. Magreñán. Extending the applicability of Gauss-Newton method for convex composite optimization on Riemannian manifolds. *Appl. Math. Comput.*, 249:453–467, 2014.
- [8] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.
- [9] J.-P. Dedieu, P. Priouret, and G. Malajovich. Newton’s method on Riemannian manifolds: convariant alpha theory. *IMA J. Numer. Anal.*, 23(3):395–419, 2003.
- [10] J.-P. Dedieu and M. Shub. Multihomogeneous Newton methods. *Math. Comp.*, 69(231):1071–1098 (electronic), 2000.
- [11] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.
- [12] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353, 1999.
- [13] O. P. Ferreira, M. L. N. Gonçalves, and P. R. Oliveira. Convergence of the Gauss-Newton method for convex composite optimization under a majorant condition. *SIAM J. Optim.*, 23(3):1757–1783, 2013.
- [14] O. P. Ferreira and R. C. M. Silva. Local convergence of Newton’s method under a majorant condition in Riemannian manifolds. *IMA J. Numer. Anal.*, 32(4):1696–1713, 2012.
- [15] O. P. Ferreira and B. F. Svaiter. Kantorovich’s theorem on Newton’s method in Riemannian manifolds. *J. Complexity*, 18(1):304–329, 2002.
- [16] O. P. Ferreira and B. F. Svaiter. Kantorovich’s majorants principle for Newton’s method. *Comput. Optim. Appl.*, 42(2):213–229, 2009.
- [17] O. P. Ferreira and B. F. Svaiter. A robust Kantorovich’s theorem on the inexact Newton method with relative residual error tolerance. *J. Complexity*, 28(3):346–363, 2012.
- [18] D. Gabay. Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.*, 37(2):177–219, 1982.
- [19] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.
- [20] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms: Fundamentals. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.
- [21] N. Karmarkar. Riemannian geometry underlying interior-point methods for linear programming. In *Mathematical developments arising from linear programming (Brunswick, ME, 1988)*, volume 114 of *Contemp. Math.*, pages 51–75. Amer. Math. Soc., Providence, RI, 1990.
- [22] S. G. Krantz and H. R. Parks. *The implicit function theorem*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. History, theory, and applications, Reprint of the 2003 edition.

- [23] S. Lang. *Differential and Riemannian manifolds*, volume 160 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 1995.
- [24] C. Li and J. Wang. Newton’s method on Riemannian manifolds: Smale’s point estimate theory under the γ -condition. *IMA J. Numer. Anal.*, 26(2):228–251, 2006.
- [25] C. Li and J. Wang. Newton’s method for sections on Riemannian manifolds: generalized covariant α -theory. *J. Complexity*, 24(3):423–451, 2008.
- [26] C. Li, J.-H. Wang, and J.-P. Dedieu. Smale’s point estimate theory for Newton’s method on Lie groups. *J. Complexity*, 25(2):128–151, 2009.
- [27] J. H. Manton. A framework for generalising the Newton method and other iterative methods from Euclidean space to manifolds. *Numer. Math.*, 129(1):91–125, 2015.
- [28] S. A. Miller and J. Malick. Newton methods for nonsmooth convex minimization: connections among U-Lagrangian, Riemannian Newton and SQP methods. *Math. Program.*, 104(2-3, Ser. B):609–633, 2005.
- [29] J. Moser. A new technique for the construction of solutions of nonlinear differential equations. *Proc. Nat. Acad. Sci. U.S.A.*, 47:1824–1831, 1961.
- [30] J. Nash. The imbedding problem for Riemannian manifolds. *Ann. of Math. (2)*, 63:20–63, 1956.
- [31] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.
- [32] Y. E. Nesterov and M. J. Todd. On the Riemannian geometry defined by self-concordant barriers and interior-point methods. *Found. Comput. Math.*, 2(4):333–361, 2002.
- [33] B. Owren and B. Welfert. The Newton iteration on Lie groups. *BIT*, 40(1):121–145, 2000.
- [34] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM J. Optim.*, 22(2):596–627, 2012.
- [35] V. H. Schulz. A Riemannian view on shape optimization. *Found. Comput. Math.*, 14(3):483–501, 2014.
- [36] M. Shub. Some remarks on dynamical systems and numerical analysis. In *Dynamical systems and partial differential equations (Caracas, 1984)*, pages 69–91. Univ. Simon Bolivar, Caracas, 1986.
- [37] S. Smale. Newton’s method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.
- [38] S. T. Smith. Optimization techniques on Riemannian manifolds. In *Hamiltonian and gradient flows, algorithms and control*, volume 3 of *Fields Inst. Commun.*, pages 113–136. Amer. Math. Soc., Providence, RI, 1994.
- [39] C. Udriste. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994.

- [40] J. H. Wang. Convergence of Newton's method for sections on Riemannian manifolds. *J. Optim. Theory Appl.*, 148(1):125–145, 2011.
- [41] J.-H. Wang, S. Huang, and C. Li. Extended Newton's method for mappings on Riemannian manifolds with values in a cone. *Taiwanese J. Math.*, 13(2B):633–656, 2009.
- [42] J.-H. Wang, J.-C. Yao, and C. Li. Gauss-Newton method for convex composite optimizations on Riemannian manifolds. *J. Global Optim.*, 53(1):5–28, 2012.
- [43] C. E. Wayne. An introduction to KAM theory. In *Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994)*, volume 31 of *Lectures in Appl. Math.*, pages 3–29. Amer. Math. Soc., Providence, RI, 1996.
- [44] Z. Wen and W. Yin. A feasible method for optimization with orthogonality constraints. *Math. Program.*, 142(1-2, Ser. A):397–434, 2013.
- [45] P. P. Zabrejko and D. F. Nguen. The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates. *Numer. Funct. Anal. Optim.*, 9(5-6):671–684, 1987.
- [46] L.-H. Zhang. Riemannian Newton method for the multivariate eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 31(5):2972–2996, 2010.